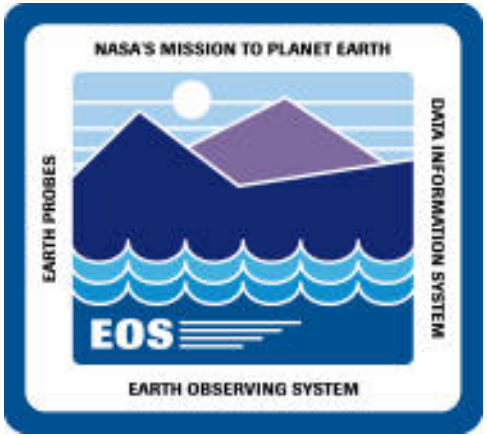


Collection Evolution

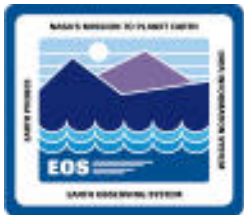
An End-to-End Scenario

Steve Marley

smarley@eos.hitc.com

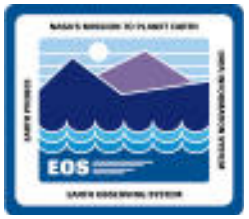
23 April 1996

Collection Evolution Scenario Overview



- Science User Scenario Description
- Functional Space Overview
- System Overview
- Pull Side Data Interaction
  - What Science & System Data is the scenario interacting with
- Push Side Data Interaction
  - How did the Science Data become available

User Scenario Description (1)



Basis:

- **Science User Scenario #7 (Dan Baldwin, U of Col)**
- **Reference: ECS User Characterization Methodology and Results, 194-00313TPW, Sept. 1994**
- **Modified to emphasize Release B data sets**

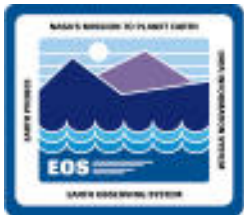
Motivation:

- **User will use coincident radiances from multiple instruments in a data fusion analysis**
- **User has preference for a specific 512 x 512 km region in western U.S.**

Notes:

- **ASTER Level 2 products are produced on demand at EDC**
- **Landsat 7 Level 0R products are maintained at EDC**
- **MODIS Level 1 products are maintained at GSFC**
- **MODIS Level 2 (Land) products are maintained at EDC**

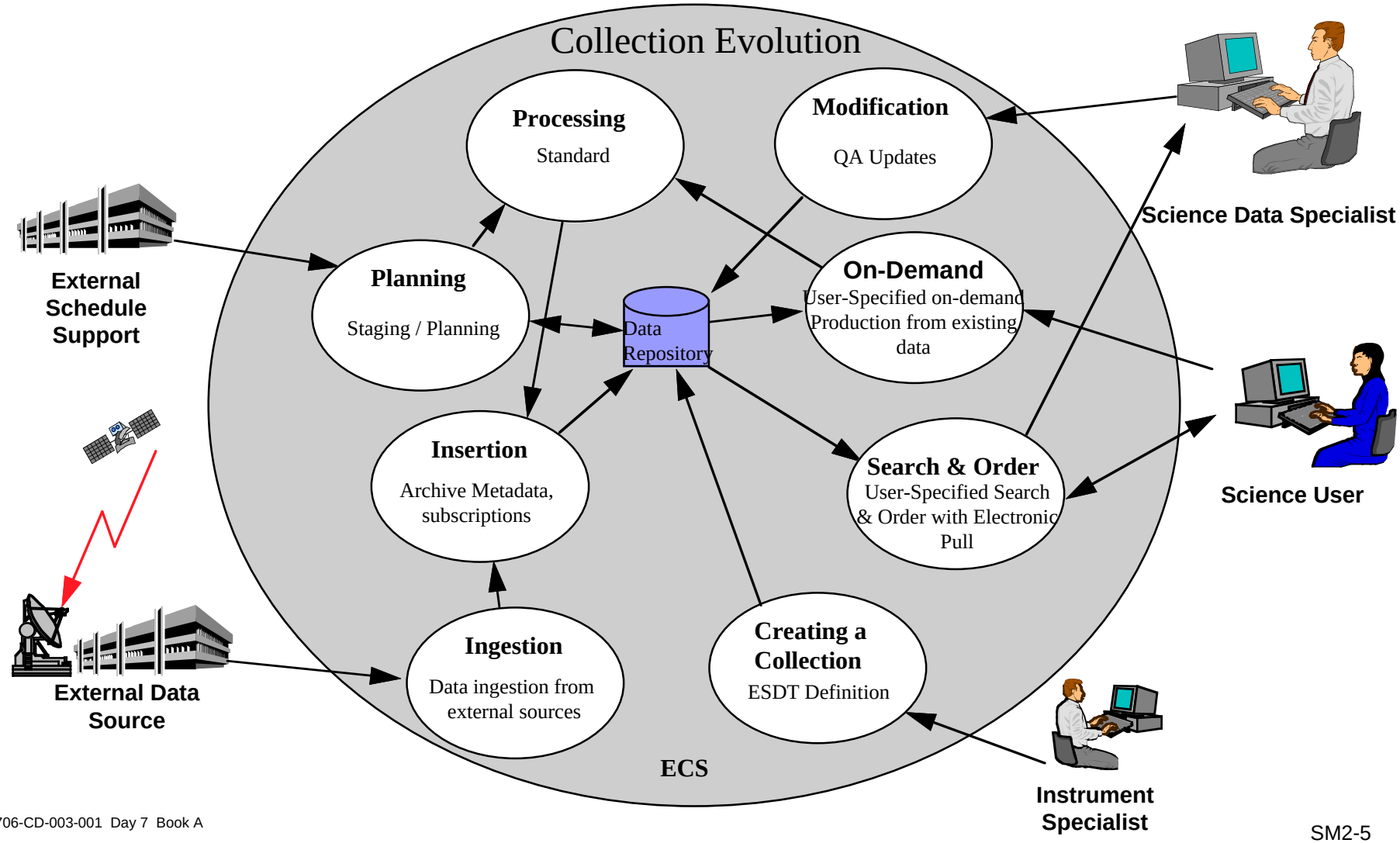
User Scenario Description (2)



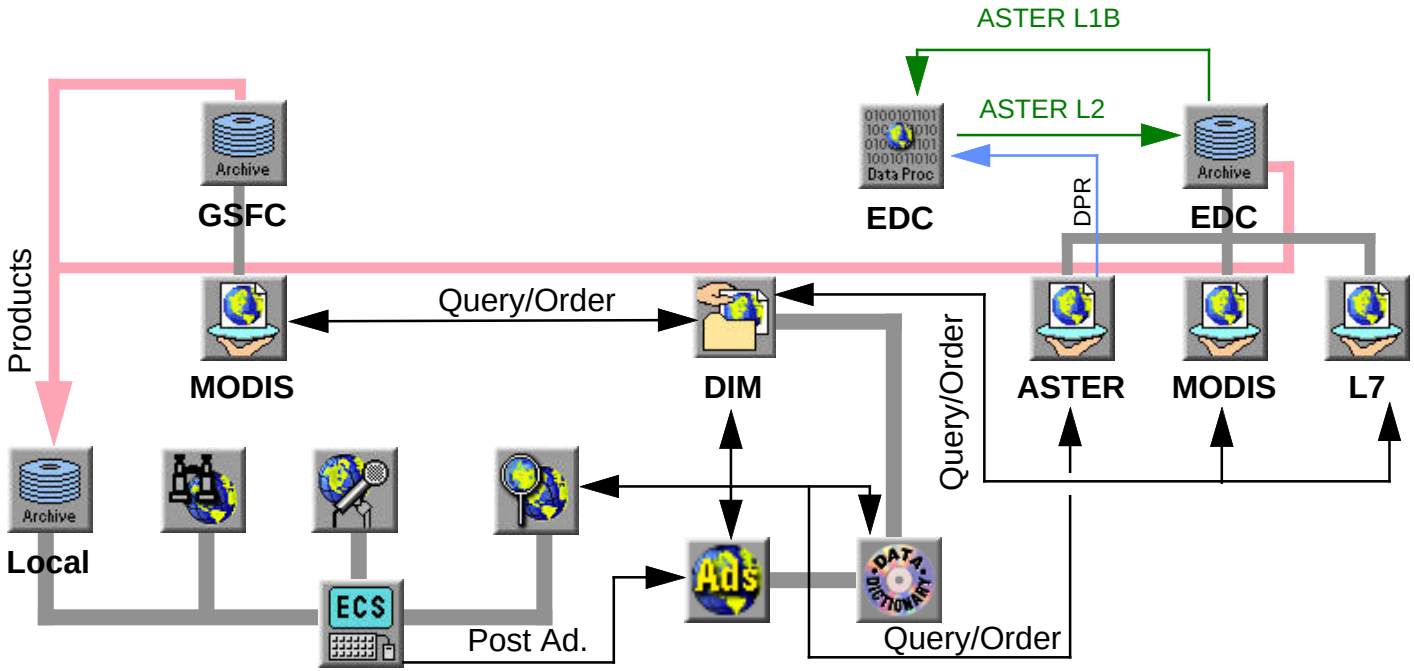
Steps (from user perspective):

1. **Search & review ASTER granules within region of interest -**
   a. **Review L2 browse images for granules which have been processed to create surface radiance products (VNIR/SWIR and/or TIR)**
   b. **Review L1A browse images for granules which have not been processed to create surface radiance products**
2. **Identify and request specific ASTER L2 granules -**
   a. **Select several cloud-free granules with appropriate surface features**
   b. **Request L2 processing to create surface radiances for granules which have not been processed with desired run-time parameters**
3. **Search for Landsat 7 L0R and MODIS L2 data coincident (+/- 45 days) with selected ASTER scenes. Review Landsat 7 and MODIS browse images to find cloud-free views.**
4. **Request Landsat 7 L0R and MODIS L1B and L2 surface radiance products subsetted to match ASTER scenes.**
5. **Analyze the delivered data at user site. Create a special collection of selected granules.**

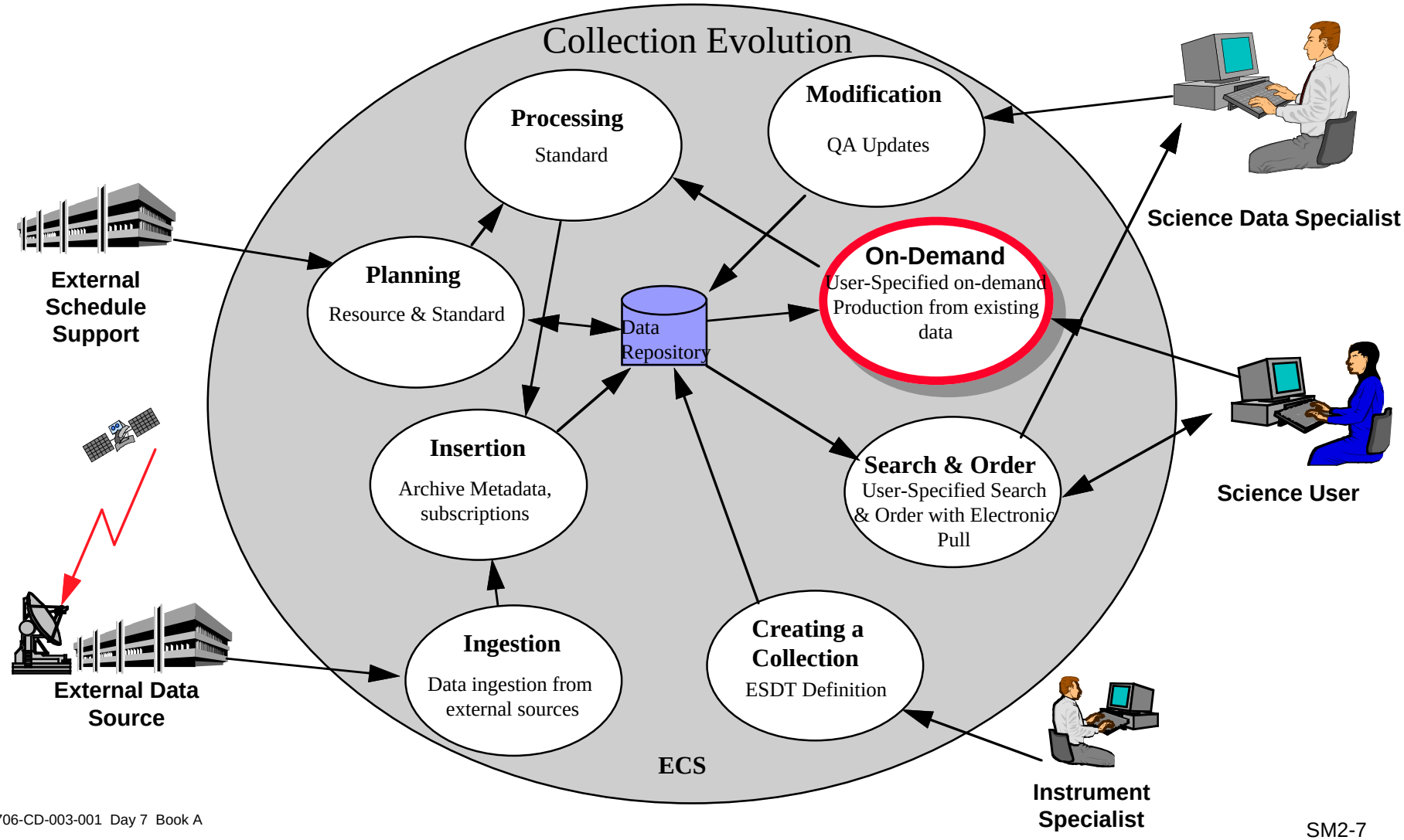
Functional Space Overview



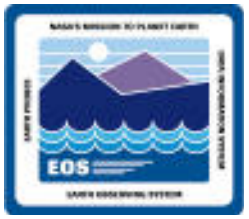
System View of Scenario



Functional Space - On-Demand Processing

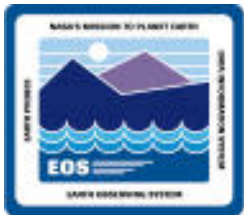


DPR Generation and Processing Process Description I



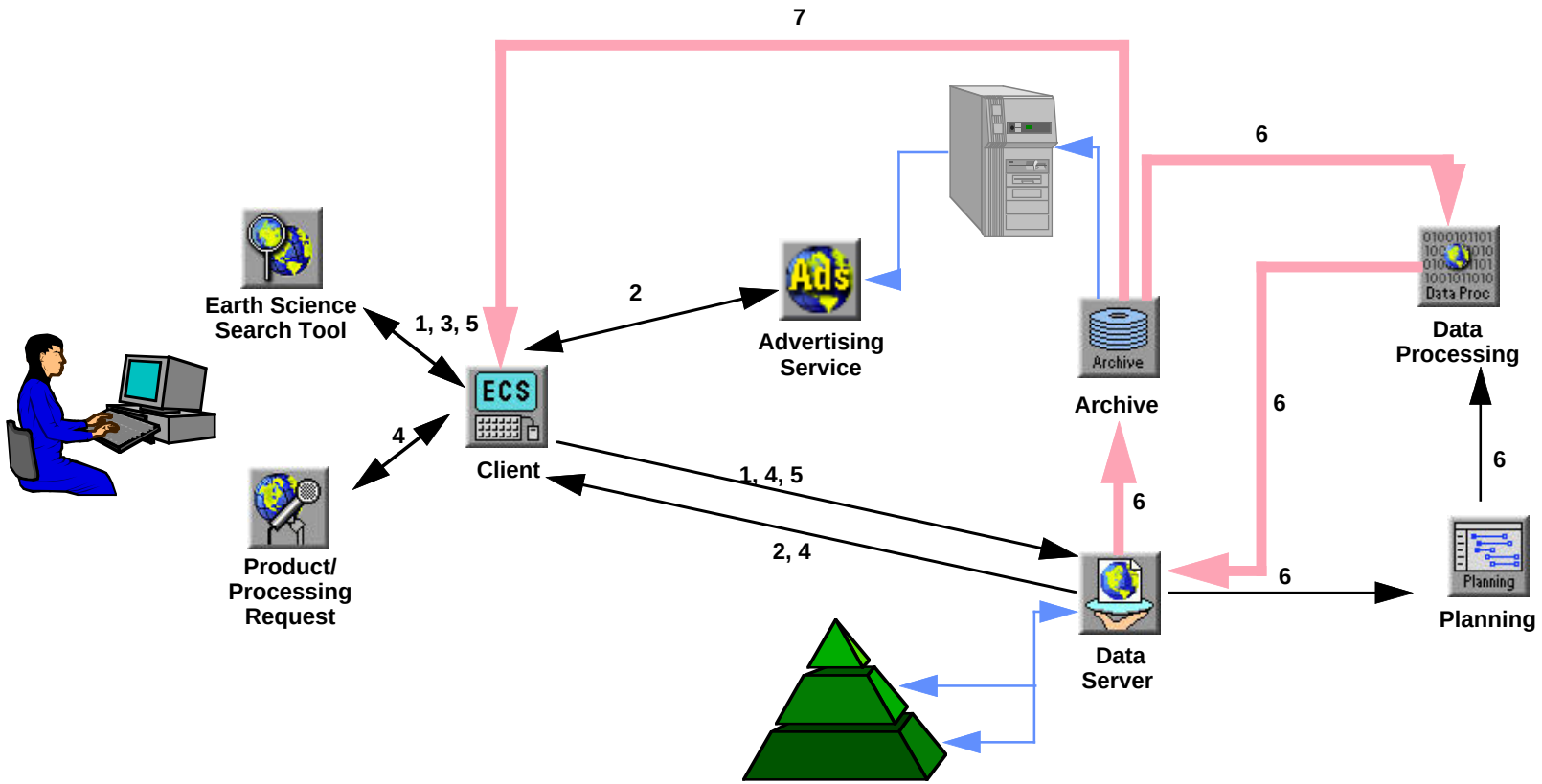
1. **Initiate Query for ASTER**
   - **User submits a query for ASTER Surface Emissivity. The query is routed to the ASTER data Server at EDC. As part of the query decomposition, the ESST formulates sub-queries for ASTER L2 & ASTER L1B consistent with the query constraints supplied by the user.**
2. **Results Notification**
   - **Upon completion of the query the Data Server notifies the ESST and sends it the UR of the results set. The ESST then accesses the result set, and displays both existing ASTER Level 2 Surface Emissivity Products and all ASTER Level 1B granules which satisfy the query. Concurrently, the ESST queries the Advertising Service for on-demand production services available for ASTER L1B, to establish the availability of these options.**
3. **Display of Processing Options**
   - **The User selects to view processing options for an ASTER L1B granule.**

DPR Generation and Processing Process Description II

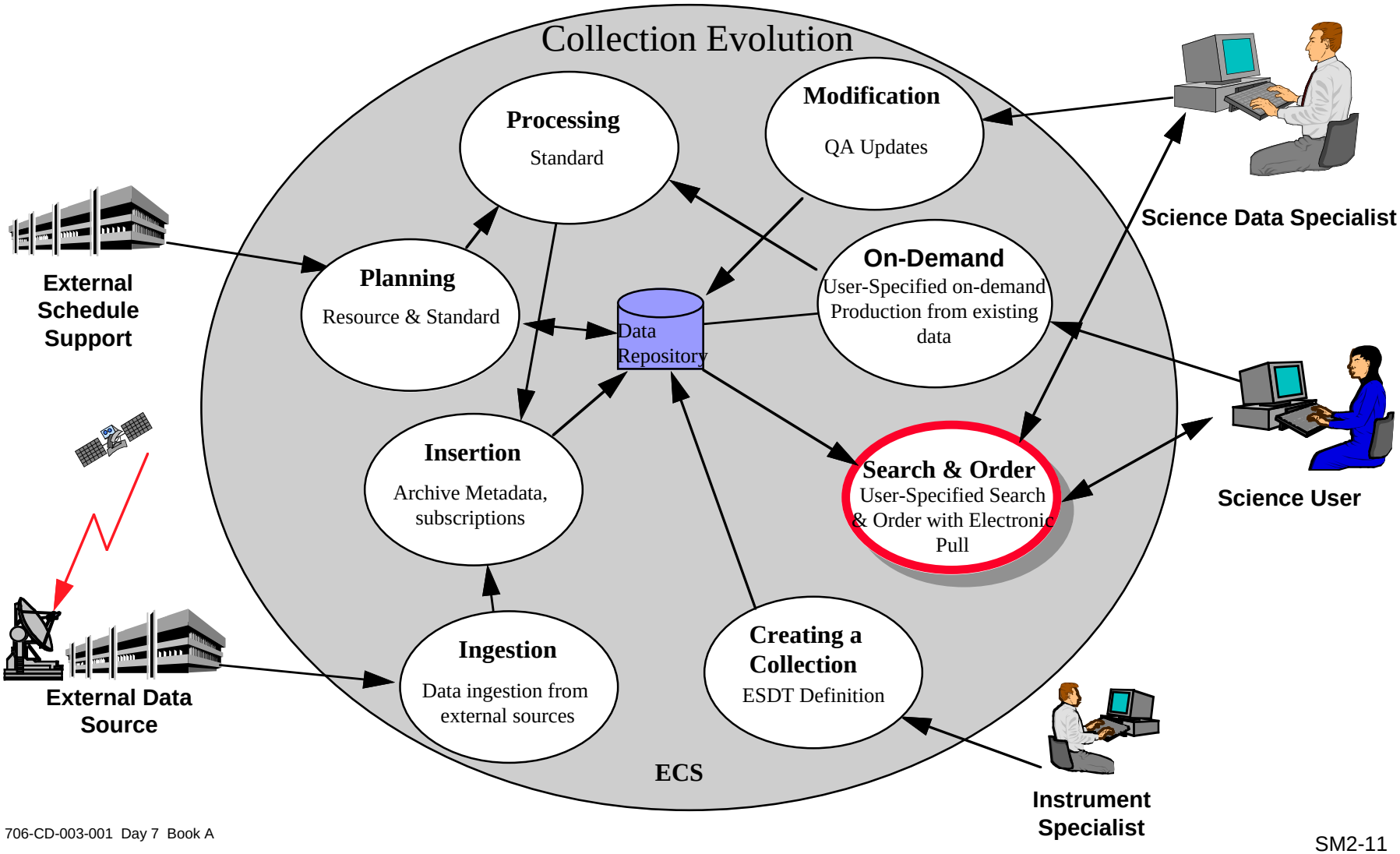


4. Data Product Customization
   - The User elects to request a ASTER Surface Emissivity without Fine Cloud Classifier. Dependency products and run-time parameters are selected. Available ancillary data sets are identified and selected. Finally the completed product order is transferred to the ESST Product Order tool, and added to other granule requests.
5. Product Ordering
   - The ESST then submits an order for the desired granules. The product order is then parsed by the Data Server. Archived granules are obtained from the archive and made ready for distribution to the user. DPR requests are routed to PDPS.
6. Producing the On-Demand Product
   - Production is initiated. Resulting granules are sent to the EDC data server for storage.
7. Order Completion
   - The granules from the DPR are assembled along with any other granules in the data order for distribution to the users. The data server notifies the user of the completion of the data order. Data is distributed to the user as specified in the original data order.

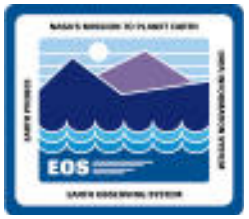
DPR Generation System Description



Functional Space-Search & Order



Coincident Search & Order Process Description I

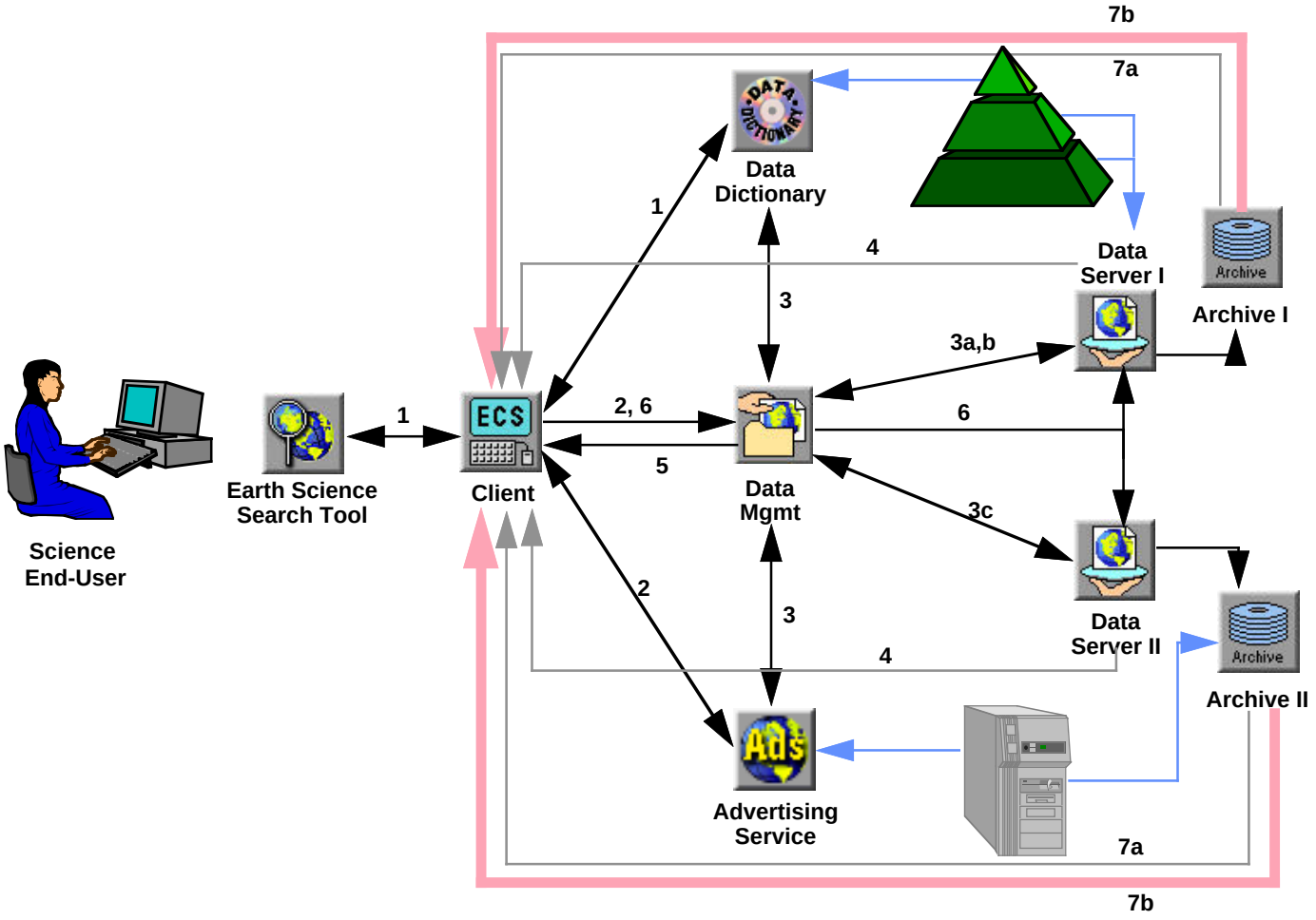


1. **Initiate ESST**
   - ESST is initiated on the Client. A request for Dictionary Valids is made. Upon the return of the valids set, the ESST menus and dialog boxes are configured.
2. **Initiate Query**
   - User submits the distributed coincident query. ESST uses Dictionary Valids to determine query routing. The Advertising Service is queried for the UR for the DIM, and the query is then dispatched.
3. **Distributing the Query**
   - The DIM initiates its query plan. It then queries the Dictionary Valids to determine the query routing, and queries the Advertising Service for the URs of the Data Servers. Coincident searches are implemented using a 'seed-set' approach
     a. The DIM queries the 'seed-set' Data Server.
     b. Upon receiving the result set from the seed query,
     c. The DIM distributes the decomposed query to the remaining Data Servers and collates the results.
4. **Performing the Query**
   - The Data Servers perform the queries, while status is monitored by the DIM and returned to the ESST. The result sets from each Data Server are returned to the DIM upon completion.

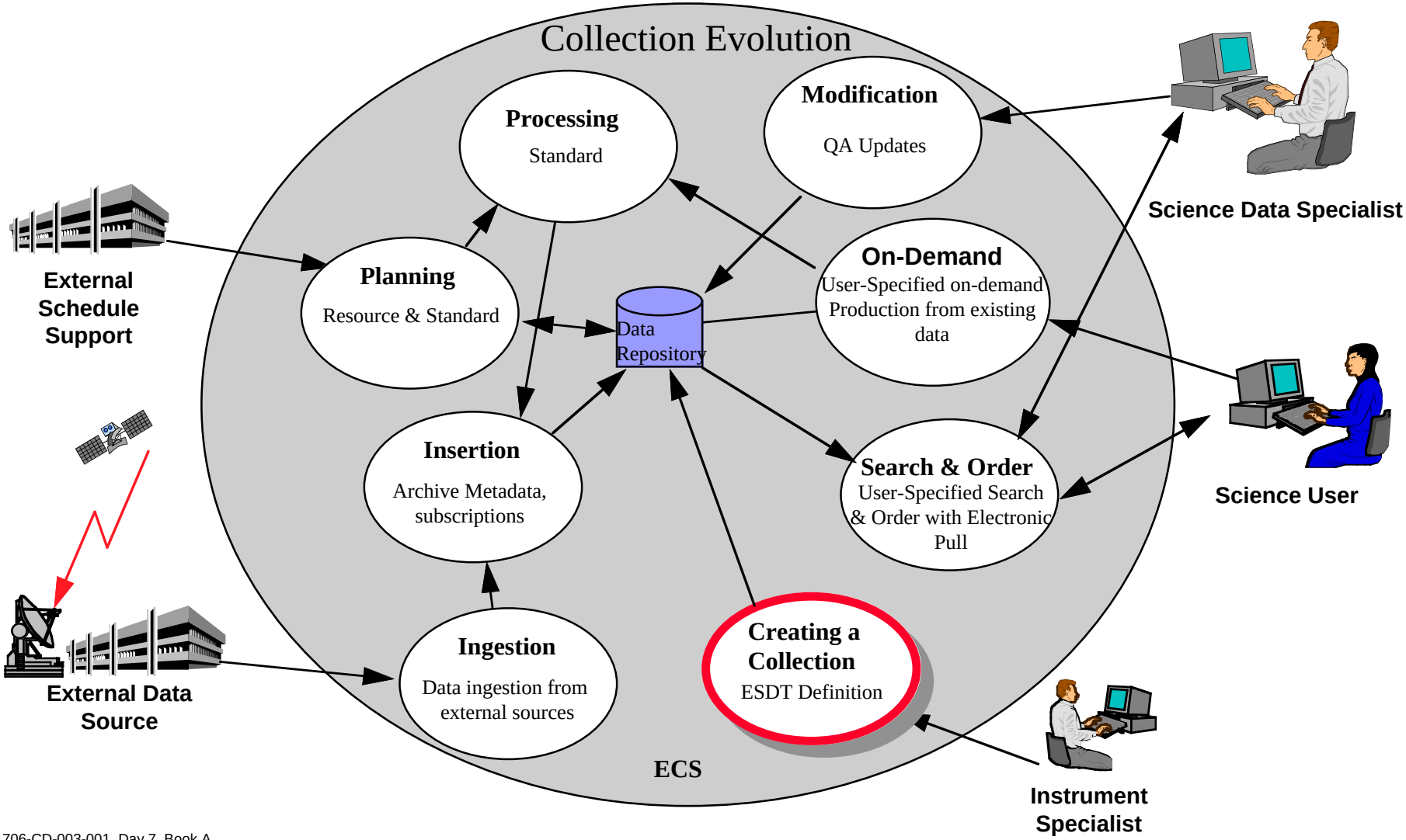
Coincident Search & Order Process Description II

5. **Results Notification**
   - Upon completion of the query the DIM notifies the ESST that the results set is complete. The ESST then accesses the result set.
6. **Product Ordering**
   - The ESST then submits an order for the desired granules to the DIM, which places orders at the appropriate Data Servers.
7. **Data Distribution**
   a. The requested data is then retrieved from the archive and the user is notified of when the data is available and where.
   b. User invokes an ftp ‘get’ to retrieve the file

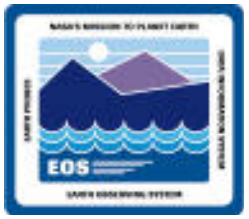
Coincident Search & Order System Description



Functional Space - A New Collection



Building a New Collection Process Description I



1. **Define & Build ESDTs**
   - **Scientists define the ESDTs both data structures and services. This information is passed to the DAAC. From this information the Data Specialist at the DAAC generates the DLLs & Descriptors necessary to instantiate the ESDTs.**
2. **Test the New ESDTs**
   - **The new ESDTs are configured on the test Data Server. A restricted access collection is created on the test Data Server complete with test data sets. Both Scientists, and DAAC staff access the collection to evaluate the new ESDT.**

Building a New Collection Process Description II

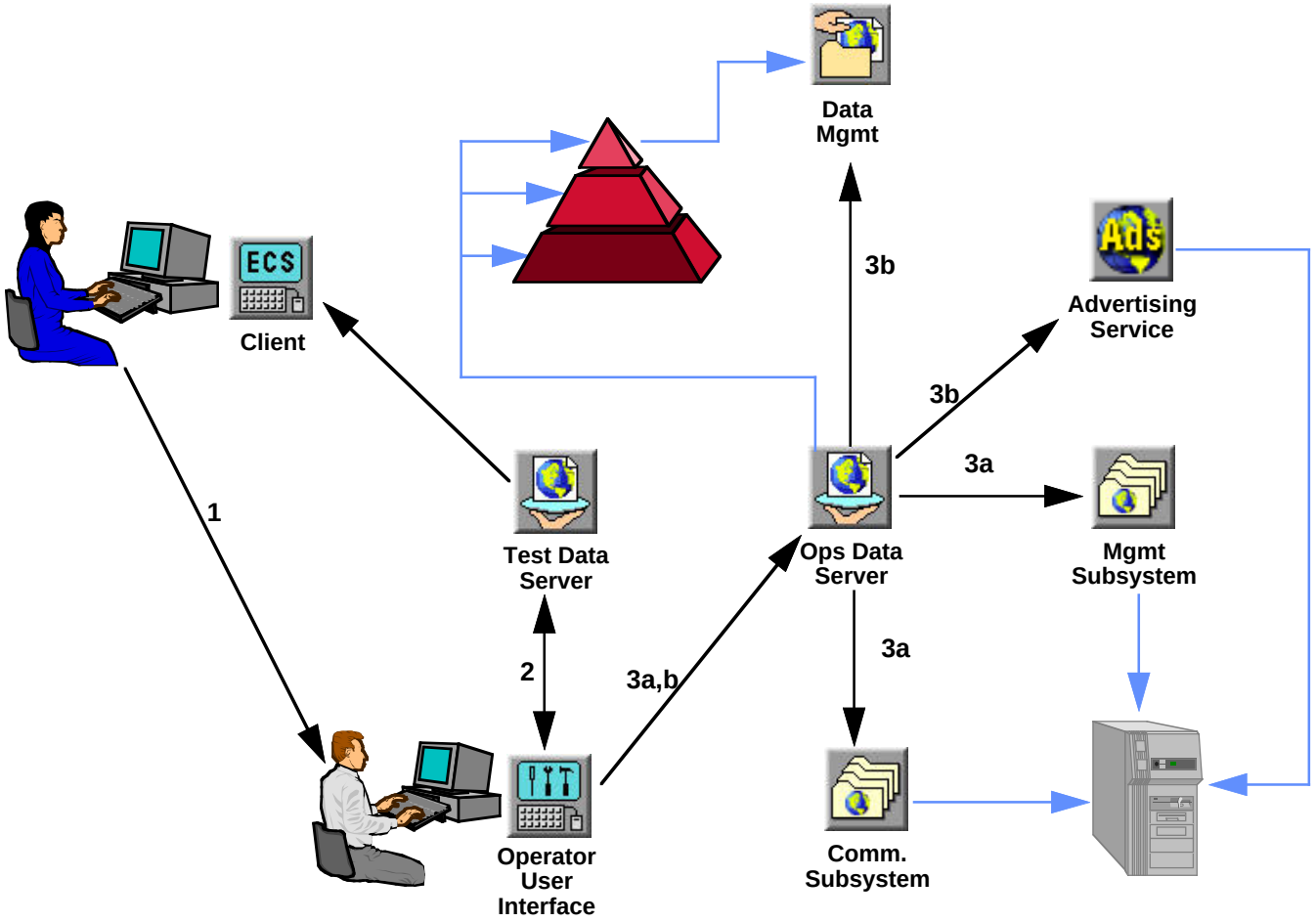
3. Install the New Collection

   a. Having approved the implementation of the ESDTs, it is then installed on the operational data server. This involves a number of steps.

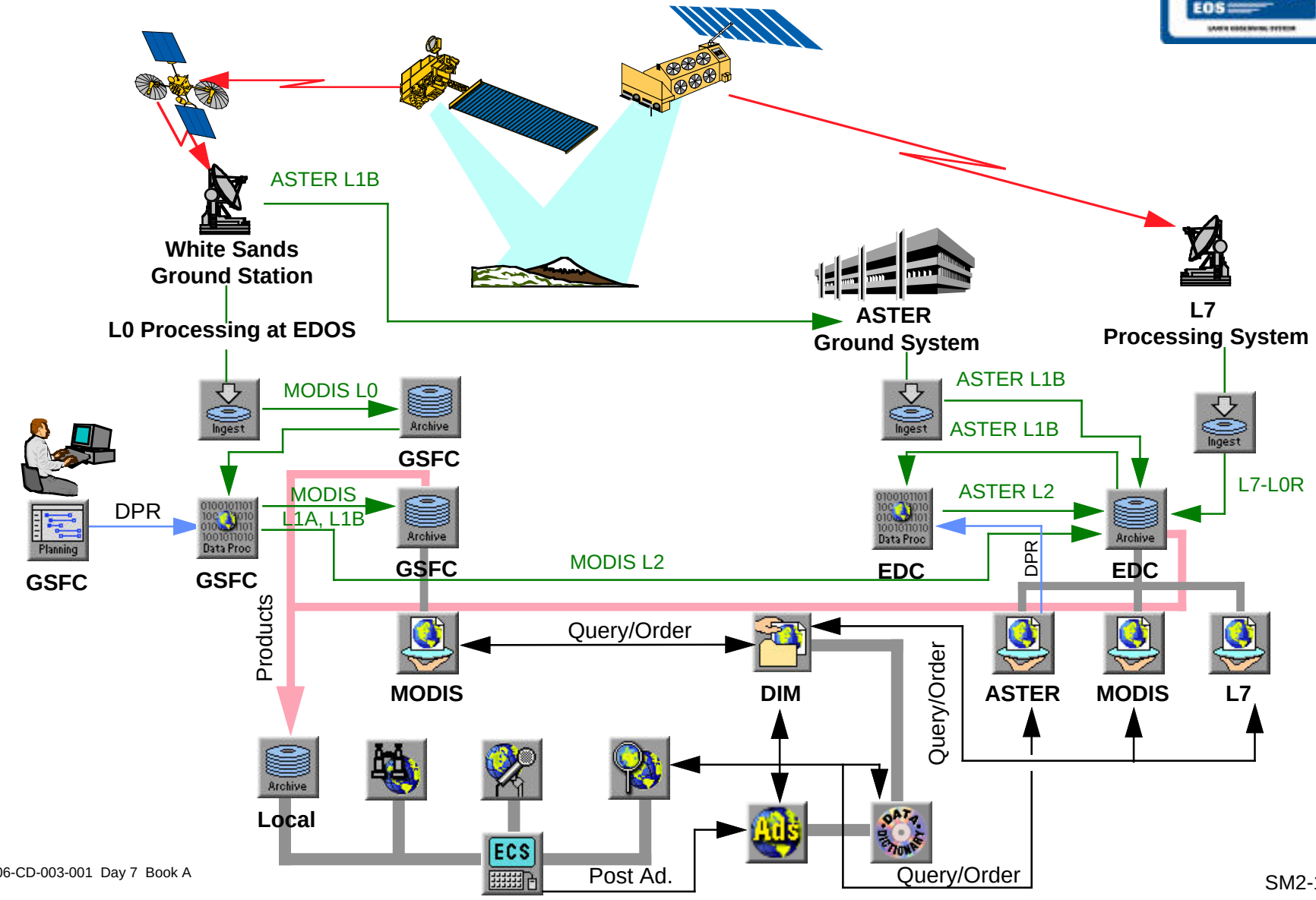
      - The ESDTs Must be registered with CSS as known resources.
      - The access controls must be established within the Data Server
      - An account entry has to registered with MSS, and
      - Subscribable events must be set up.

   b. Once the collection is defined in the Data Server, then the relevant collection level metadata is exported to Data Management, and an advertisement posted. The collection is then available across ECS.

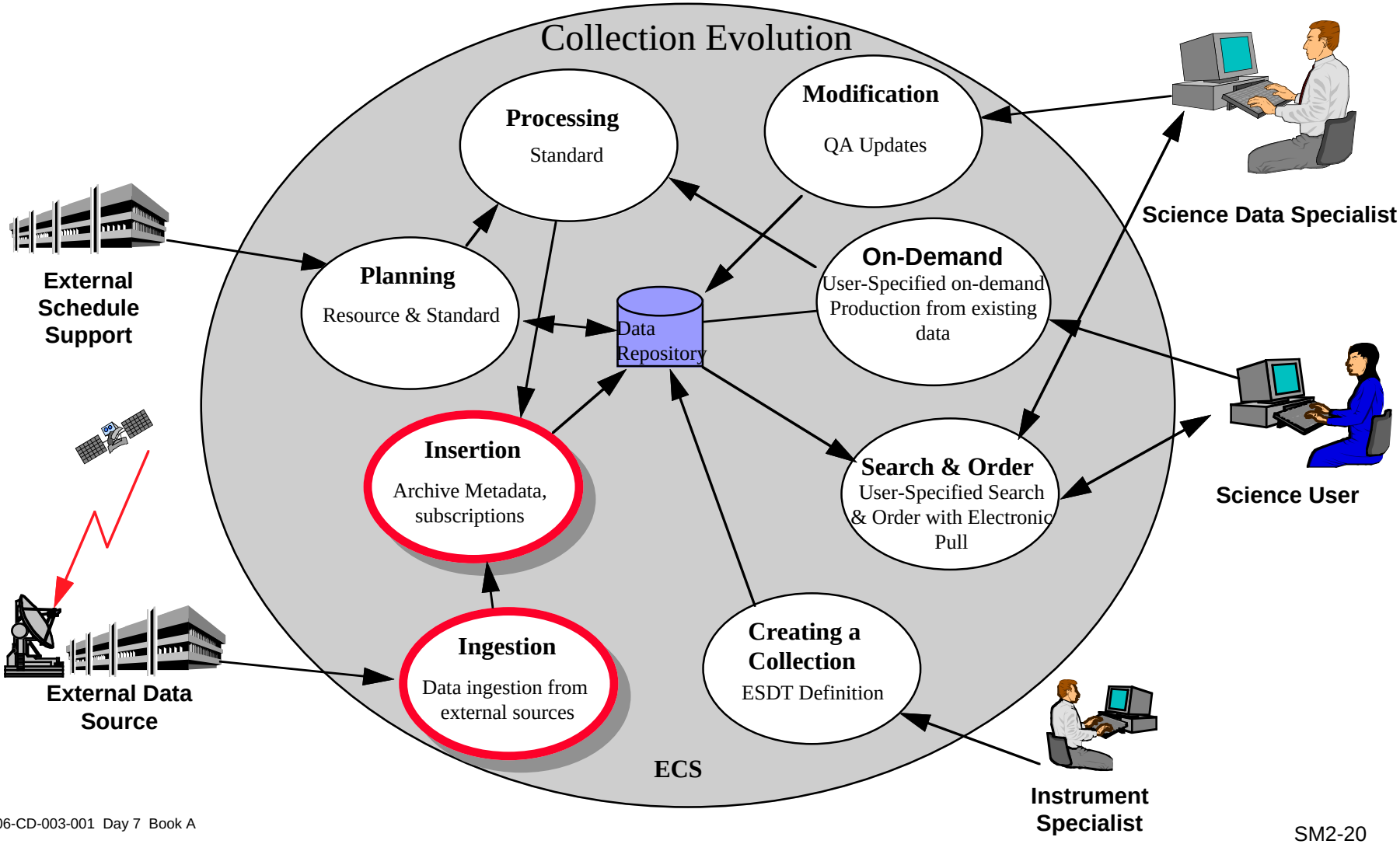
Building a New Collection System Description



Complete Overview



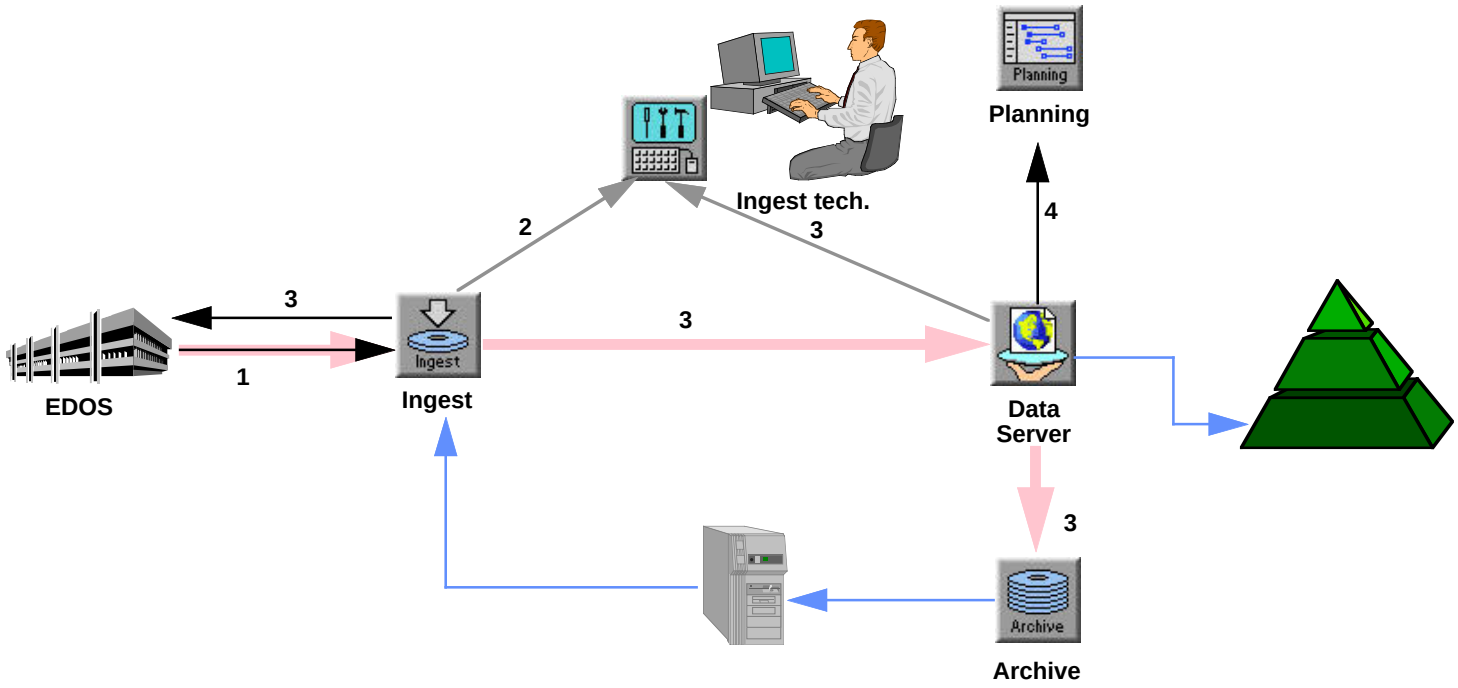
Functional Space - Ingest & Insert



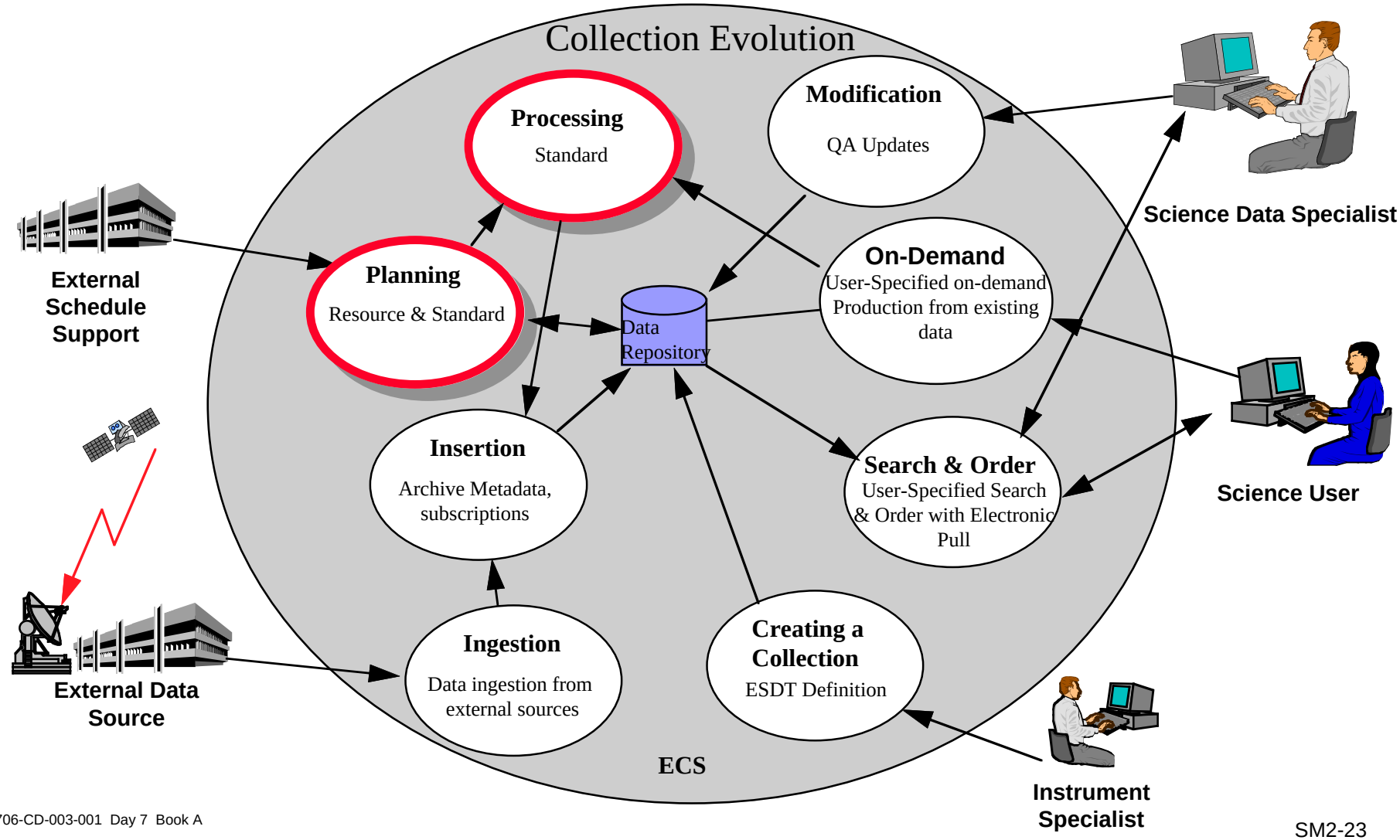
L0 Data Ingest & Insert Process Description

1. **Data Transfer**
   - The Production Data Set (PDS) is transferred to the ECS disk space. The PDS delivery record is sent to a monitored directory. This directory is regularly polled by Ingest and the PDS Delivery Record is detected.
2. **Prepare Data for Archive**
   - Ingest interprets the Delivery Record, and accesses the L0 data file. The file is then checked against the Delivery Record for required metadata, and metadata valid ranges. Any errors are reported to the Ingest technician.
3. **Archive the Data**
   - Ingest then generates an insert command, and attempts to insert the data to the L0 archive. Upon successful insert a notification is returned to EDOS. Any errors are reported to the Ingest technician.
4. **Subscription Servicing**
   - Data insert is one of the events that is recognized by Subscription Services. Upon insert of data into the archive Subscription Services registers subscriptions against that data set (in this case principally Planning), and issues a notification.

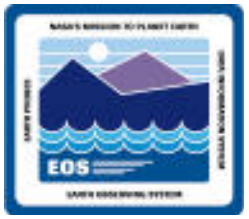
L0 Ingest & Insert System Description



Functional Space-Planning & Processing

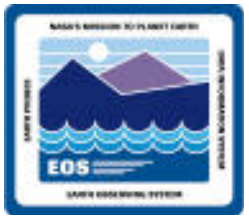


Planning & Processing Process Description I



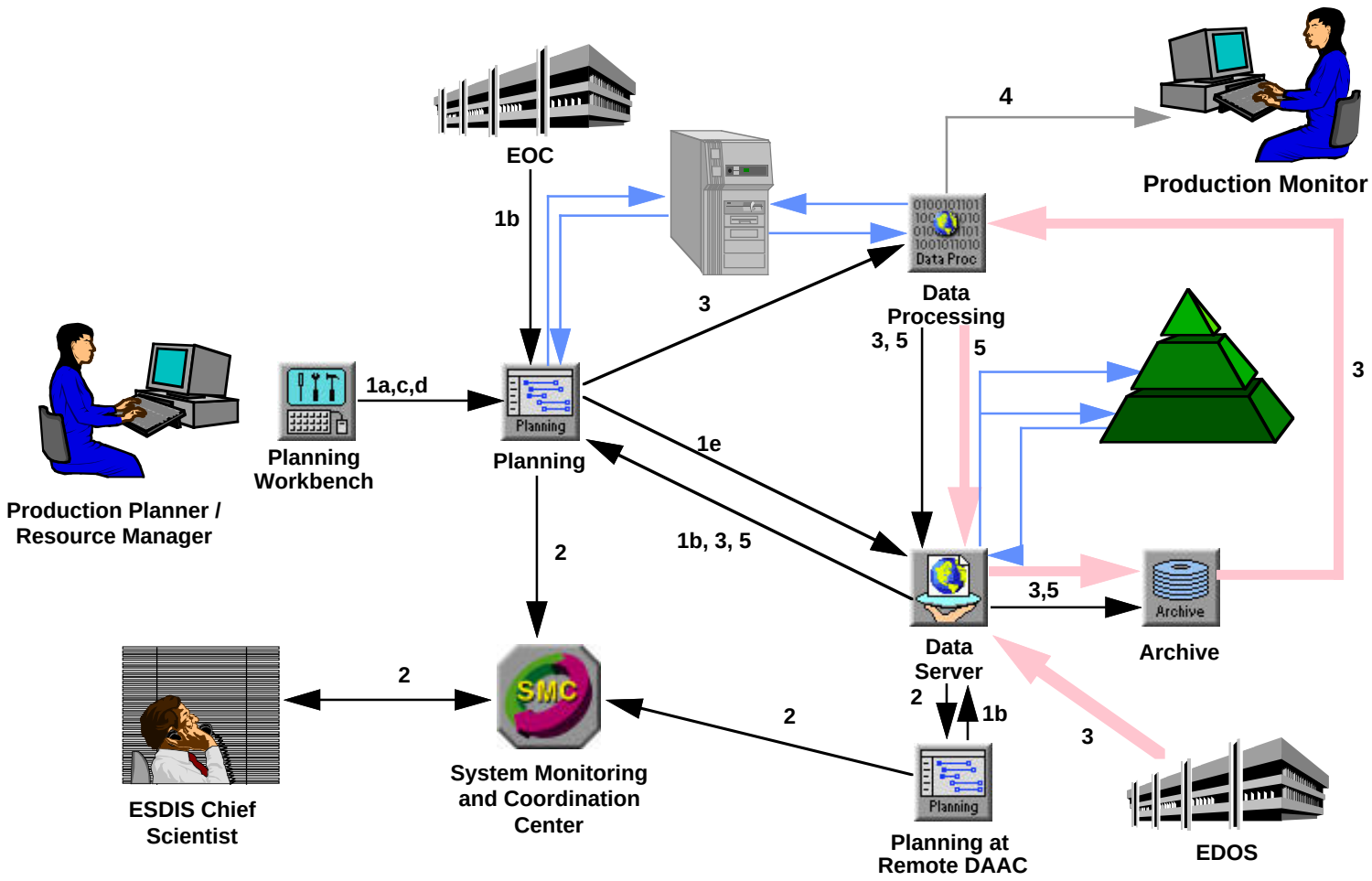
1. Planning at a Single Site
   a. Approved resource request are entered into the Planning data base, and a new resource plan is generated.
   b. The latest Data Activity Schedule for AM-1 is incorporated, together with the latest Data Availability Schedules from remote DAACs.
   c. Reprocessing and on-demand thresholds, production strategies and Limited Automatic and Replan thresholds are defined.
   d. Routine and Reprocessing Production Request are defined
   e. Planning is initiated. The new plan is archived to the Data Server and entered into the Processing Queue
2. Cross-DAAC Planning
   - Receive new candidate plan from another DAAC. Compare the plan with the baseline, and communicate any recommended changes back to original DAAC. SMC provides role-up information to ESDIS Project Scientist.

Planning & Processing Process Description II

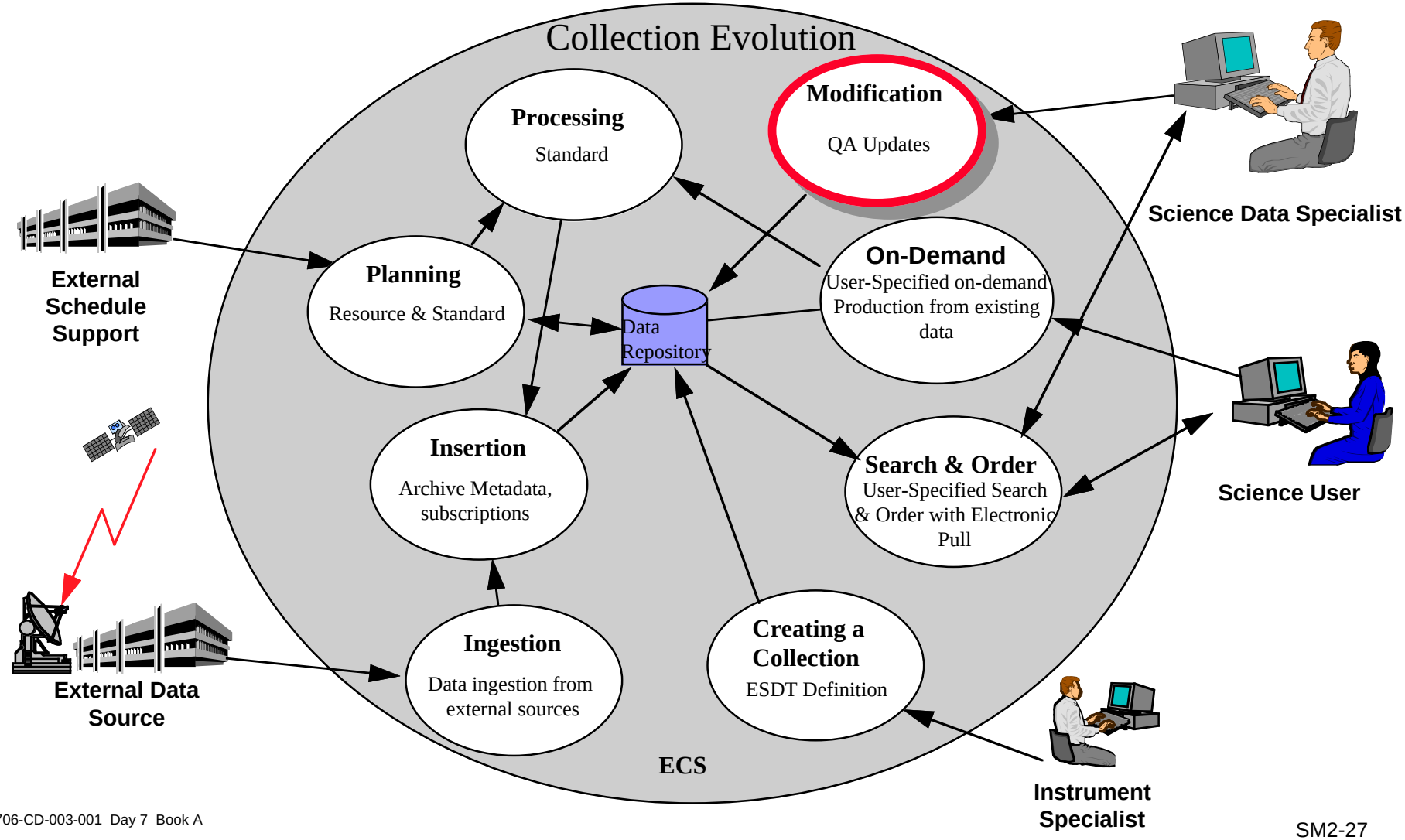


3. Standard Production
   - L0 data arrives from EDOS. Planning is notified of data availability, and releases jobs for processing. Data is staged from the Archive, and processing commences.
4. Anomalous Exit conditions
   - If processing completes with an anomalous condition, the Production Monitor is notified, and predefined actions are invoked. These could include:
     a. running a QA job,
     b. resubmitting the job, or
     c. simply sending the failure notice to the SCF.
5. Successful completion
   - Output products are destaged from the Processing hardware, and inserted into the Data server.

Planning & Processing System Description



Functional Space - QA



Quality Assurance Process Description I

1. In-line PGE QA
   - Data Processing generates a product. In-line QA is performed at the end of the run and PGE QA flags set accordingly. The Production Monitor is notified of completion status.
2. Data Insertion/Visibility
   - The product is inserted into the archive, and if the QA flags are correctly set the product is made visible to the general end-user, else it will only be seen by privileged users. Appropriate subscriptions are triggered and users are notified.
3. Off-line DAAC QA
   - The DAAC Data Specialist receives notification of a new product and retrieves the product. Off -line QA is then performed.
4. QA Update
   - The DAAC Data Specialist submits a QA update against the product. This may trigger a change in the visibility of the product depending upon the nature of the update.

Quality Assurance Process Description II

5. Off-line SCF QA and Update

- **The SCF has the same capability as the DAAC staff for updating the QA of a product through success to SCF specific QA flagging at the granule level. The final accessibility of a product to the end-user, and/or internal user (e.g. Planning) is configurable based on predefined rules for each product type (ESDT).**

Quality Assurance System Description